

Lossy Source Coding via Spatially Coupled LDGM Ensembles

Vahid Aref, Nicolas Macris, Rüdiger Urbanke and Marc Vuffray
LTHC, IC, EPFL, Lausanne, Switzerland,

Abstract—We study a new encoding scheme for lossy source compression based on spatially coupled low-density generator-matrix codes. We develop a belief-propagation guided-decimation algorithm, and show that this algorithm allows to approach the optimal distortion of spatially coupled ensembles. Moreover, using the survey propagation formalism, we also observe that the optimal distortions of the spatially coupled and individual code ensembles are the same. Since regular low-density generator-matrix codes are known to achieve the Shannon rate-distortion bound under optimal encoding as the degrees grow, our results suggest that spatial coupling can be used to reach the rate-distortion bound, under a *low complexity* belief-propagation guided-decimation algorithm.

This problem is analogous to the MAX-XORSAT problem in computer science.

Index Terms—Lossy source coding, spatial coupling, LDGM, belief propagation guided decimation, rate distortion bound.

I. INTRODUCTION

The spatial coupling of copies of a graphical code was introduced in [1] in the form of convolutional low-density parity-check (LDPC) codes. The performance of such ensembles under the belief propagation (BP) algorithm is consistently better than the performance of the underlying ensembles [2–4]. The key observation is that the BP threshold of a coupled ensemble considerably improves and gets close to the maximum a posteriori (MAP) threshold of the underlying ensemble. This *threshold saturation* phenomenon has been studied rigorously in [5, 6]. Furthermore, it has been investigated in other models such as the Curie-Weiss chain, random satisfiability, graph coloring [7–9], and compressed sensing [10, 11].

One of the classic problems in communications is lossy source compression. The objective is to compress a given sequence, so that it can be reconstructed up to some specified distortion. For binary symmetric sources, low-density generator-matrix (LDGM) codes are able to asymptotically achieve Shannon’s rate-distortion bound under optimal encoding (minimum distance encoding) [12, 13]. However, this is not a computationally efficient scheme.

LDGM codes are well-suited to low-complexity message passing algorithms such as the BP algorithm. But using LDGM codes with a plain BP algorithm is not very effective in lossy compression. To achieve more promising results, one can equip the BP algorithm with a decimation process. The general idea of belief-propagation guided-decimation (BPGD) algorithms is to: i) compute BP marginals, ii) fix bits with the largest bias, iii) decimate the graph. Decimation reduces the graph to a smaller one, on which this process is repeated. Many variants of message passing algorithms and various decimation

processes have been investigated. In [14] a survey propagation (SP) inspired decimation algorithm is proposed. Simulations show distortions close to the rate-distortion bound for large block-lengths. Later, similar results were reported with modified forms of BP [15, 16]. The performance of these algorithms also depends on the choice of degree distributions for LDGM codes. A heuristic choice is to use degree distributions that have been optimized for LDPC codes on binary symmetric channel [14–16]. No rigorous analysis exists to date which can explain why this is the case.

In the present contribution, we study a spatially coupled LDGM ensemble for lossy source compression. Two of us [17] studied such ensembles in the context of rateless codes (channel coding) and demonstrated that threshold saturation takes place. We provide numerical evidence showing that a similar effect occurs in lossy compression. In particular, the BPGD distortion of the coupled LDGM ensemble approaches (numerically) the optimal distortion of the underlying ensemble. We use the simplest forms of BP and decimation processes, and we take regular degrees. It is noteworthy that no optimization is needed, thus suggesting that the observed saturation is related to fundamental principles. Regular (underlying) LDGM ensembles with large degrees have an optimal distortion that approaches the rate-distortion limit. Thus with spatially coupled LDGM codes with regular large degrees and a BPGD algorithm we can attain the Shannon limit. The complexity of the encoding scheme presented here is $O(n^2 L^2)$ where L is the number of copies of the underlying ensemble and n the size of each copy.

In section II, we briefly review lossy compression and explain the structure of coupled LDGM ensembles. In section III, we formulate lossy compression as an optimization problem and compute the optimal distortion for underlying and coupled LDGM ensembles, by using the SP formalism. In section IV, we formulate and discuss the BPGD algorithm. Simulation results for this algorithm are presented in section V. A few practical issues are discussed in section VI.

II. FRAMEWORK

1) Lossy Compression of Symmetric Bernoulli Sources:

Let $\underline{X} \in \mathcal{X} = \{0, 1\}^n$ represent a binary source of length n . We have $\underline{X} = \{X_1, X_2, \dots, X_n\}$, where $\{X_a\}_{a=1}^n$ are i.i.d random variables with $\mathbb{P}\{X_a = 1\} = \frac{1}{2}$, for $a \in \{1, \dots, n\}$. We compress a given source word \underline{x} by mapping it to one of the 2^{nR} index words $\underline{u} \in \mathcal{U} = \{0, 1\}^{nR}$, where $R \in [0, 1]$ is the rate. The stored sequence \underline{u} then defines a reconstructed source sequence $\hat{\underline{x}}(\underline{u}) \in \mathcal{X}$, where the source decoding map

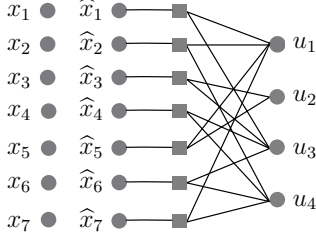


Fig. 1. The factor graph associated to an underlying LDGM code.

$\underline{u} \rightarrow \hat{\underline{x}}(\underline{u})$ depends on the structure of the code. For a given pair $(\underline{x}, \hat{\underline{x}})$, the distortion is measured by Hamming distance $\frac{1}{n}d_H(\underline{x}, \hat{\underline{x}}) = \frac{1}{n} \sum_{a=1}^n |x_a - \hat{x}_a|$. Thus, the average quality of the reconstruction is measured by $D := \frac{1}{n} \mathbb{E}_{\underline{X}}(d_H(\underline{x}, \hat{\underline{x}}))$.

For the symmetric Bernoulli source, it is well-known that for any scheme, the average distortion is lower-bounded by D_{sh} , defined implicitly by the rate-distortion function $h(D_{\text{sh}}) = 1 - R$, $D_{\text{sh}} \in [0, \frac{1}{2}]$. Here $h(\cdot)$ is the binary entropy function. The goal is to find an encoding scheme such that, for a given R , it achieves the rate-distortion lower bound.

For this purpose, we study LDGM codes since they are able to achieve the rate-distortion bound under the optimal encoding when the average degrees increase [12].

2) *LDGM Codes*: Consider an LDGM code of block length n and rate $R = \frac{m}{n}$. Let u_1, \dots, u_m be m code-bits of the index word \underline{u} and let $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n$ be n reconstruction bits of the source word x_1, x_2, \dots, x_n . An LDGM code is usually represented by a bipartite graph as depicted in Figure 1. Call this graph $\Gamma(C, G; E)$. Each code-bit u_i is represented by a node $i \in C(\Gamma)$. For each reconstruction bit \hat{x}_a , there is a *generator node* $a \in G(\Gamma)$. Each edge $(i, a) \in E(\Gamma)$ shows that the code-bit u_i is connected to the corresponding reconstruction bit \hat{x}_a . We denote by ∂a the set of all code-bit nodes connected to $a \in G$, i.e. $\partial a = \{i \in C \mid (i, a) \in E\}$. Similarly, for $i \in C$, $\partial i = \{a \in G \mid (i, a) \in E\}$. Thus, the reconstructed bit \hat{x}_a is obtained from the code-bits \underline{u} by a mod 2 sum $\hat{x}_a = \bigoplus_{i \in \partial a} u_i$.

In this paper, we focus on the (l, r, n) -regular LDGM random ensemble, where l (resp. r) is the degree of generator (resp. code-bit) nodes. We refer to [18] for the detailed construction of this ensemble.

3) *Chain of LDGM(l, r, n) Ensembles*: Consider L sets of nodes each having $m = \frac{l}{r}n$ code-bit nodes and n generator nodes (and reconstruction nodes). Locate the sets in positions 0 to $L-1$ on a circle (see Figure 2). We randomly connect each generator node in position $z \in [0, L-1]$, to l code-bit nodes from w sets in the range $[z, z+w-1]$. Eventually (for large n and m) code-bit nodes have degree r . The details of this construction are explained in [5, 17]. Note that this ensemble has the same *local* structure as the underlying LDGM(l, r) ensemble. Because of the *global* circular structure, we call it a *Closed Chain LDGM(l, r, L, w, n) ensemble (CCLDGM)*.

We wish to point out a difference between the present construction and the ones used so far in channel coding. The latter are based on open linear chains with suitable boundary conditions that help initiate the decoding process, which then propagates like a wave through the system. In the present

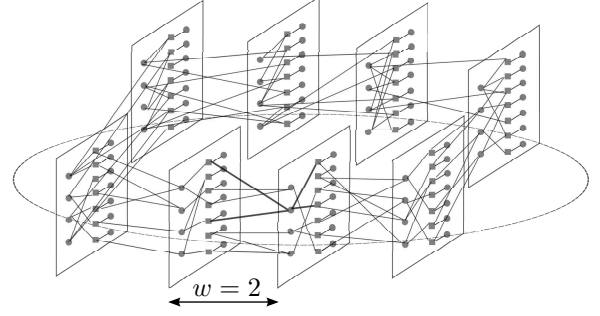


Fig. 2. Illustration of a factor graph from a CCLDGM ensemble with length $L = 8$ and $w = 2$.

periodic construction, encoding will be seeded by preferential decimation at a specific position (say $z = L/2$), in an initial phase of the BPGD algorithm.

III. OPTIMAL ENCODING AND DISTORTION

A. Optimal Encoding

Let Γ be the factor graph of a random instance of an LDGM(l, r, n) ensemble or CCLDGM(l, r, L, w, n) ensemble. We are looking for a configuration \underline{u}^* that minimizes the distortion between \underline{x} and $\hat{\underline{x}}(\underline{u})$, i.e.

$$d_{\min} := \min_{\underline{u}} \frac{1}{nL} d_H(\underline{x}, \hat{\underline{x}}) = \frac{1}{nL} d_H(\underline{x}, \bigoplus_{i \in \partial a} u_i^*). \quad (1)$$

Note that for the individual ensemble $L = 1$. As we are interested in the average distortion of the ensemble over all $\underline{X} \in \mathcal{X}$, we define the optimal distortion as

$$D_{\text{opt}} = \mathbb{E}_{\Gamma, \underline{X}}(d_{\min}). \quad (2)$$

We transform the above minimization problem into a maximum likelihood problem by equipping the configuration space $\{0, 1\}^{nL/r}$ with the following probability measure,

$$\begin{aligned} \mu_{\beta}(\underline{u} \mid \underline{x}) &:= \frac{1}{Z(\beta \mid \underline{x})} e^{-\beta d_H(\underline{x}, \hat{\underline{x}})} \\ &= \frac{1}{Z(\beta \mid \underline{x})} \prod_{a \in C(\Gamma)} e^{-\beta |x_a - \bigoplus_{i \in \partial a} u_i|}, \end{aligned} \quad (3)$$

where $\beta \in \mathbb{R}^+$ is a non-negative parameter, and $Z(\beta \mid \underline{x}) = \sum_{\underline{u}} e^{-\beta d_H(\underline{x}, \hat{\underline{x}})}$ the normalizing factor. The minimizer \underline{u}^* in (1) maximizes the measure, i.e. $\max_{\underline{u}} \mu_{\beta}(\underline{u} \mid \underline{x}) = \mu_{\beta}(\underline{u}^* \mid \underline{x})$. Moreover, for a particular \underline{x} , the minimal distortion can be obtained as $d_{\min} = -(nL)^{-1} \lim_{\beta \rightarrow \infty} \beta^{-1} \ln Z(\beta \mid \underline{x})$. This formulation of the problem allows to use techniques from statistical physics to compute D_{opt} . In the physics interpretation, d_H and μ_{β} are a random hamiltonian and Gibbs measure, β is an inverse temperature, Z a partition function, D_{opt} is the average ground state energy. The latter can be computed by the SP formalism.

B. Computation of the Optimal Distortion

The details of the SP equations, used to compute D_{opt} , are not shown here. A pedagogical account of the formalism can be found in [19]. The numerical solution of the SP equations shows that the optimal distortions for the underlying

TABLE I

THE OPTIMAL AND THE BPGD DISTORTIONS FOR THE LDGM($k, 2k$) ENSEMBLES. THE BPGD DISTORTIONS ARE EVALUATED FOR $n = 200000$ NODES. THE RATE-DISTORTION BOUND FOR $R = 0.5$ IS $D_{\text{sh}} \approx 0.1100$.

$(k, 2k)$	(3, 6)	(4, 8)	(5, 10)
D_{opt}	0.1139	0.1111	0.1105
D_{BPGD}	0.1357	0.1590	0.1811

LDGM(l, r) and CCLDGM(l, r, L, w) are the same for each finite w and L (here n and m are infinite). In fact, for an ensemble with even l and Poisson degree for code-bit nodes, a rigorous proof is presented in [9] where the spatially coupled periodic XORSAT problem is discussed. Optimal distortions for degree distributions $(k, 2k)$, $k = 3, 4$ and 5 are given in the first line of Table I. We observe that, as k increases, the optimal distortion converges to the rate-distortion bound D_{sh} . This observation is consistent with the result in [12, 13].

IV. BELIEF PROPAGATION GUIDED DECIMATION

A. Belief Propagation Approximation

The naive way to compute the partition function involves the summation over $2^{nLl/r}$ terms. Unfortunately, this approach is too complex. The same problem occurs for computing the marginal distribution of a node. On a tree, however, these computations can be done exactly and yield the BP equations. We deal with graphs that are locally tree-like and this motivates the use of the BP equations as a first approximation for the marginals. These involve a set of $2|E(\Gamma)|$ real valued messages, each pair being associated to an edge. The messages on the edge $(i, a) \in C(\Gamma) \times G(\Gamma)$ are denoted by $\eta_{i \rightarrow a}$ and $\hat{\eta}_{a \rightarrow i}$, and satisfy

$$\begin{aligned} \hat{\eta}_{a \rightarrow i} &= \frac{1}{\beta} \tanh^{-1} \left((-1)^{x_a} \tanh\left(\frac{\beta}{2}\right) \prod_{j \in \partial a \setminus i} \tanh(\beta \eta_{j \rightarrow a}) \right) \\ \eta_{i \rightarrow a} &= \sum_{b \in \partial i \setminus a} \hat{\eta}_{b \rightarrow i}. \end{aligned} \quad (4)$$

As a side remark, note that as $\beta \rightarrow \infty$ the BP equations become the so-called Min-Sum equations. From the solutions of equations (4) one can compute the BP marginal distributions of code-bit i and generator node a , and a Bethe approximation $\mu_{\beta}^{\text{BP}}(\underline{u}|\underline{x})$ (this is not a probability measure in general) of the original measure $\mu_{\beta}(\underline{u}|\underline{x})$. We refer to [19] for more information.

Unfortunately, the number of solutions of the BP equations which lead to roughly the same distortion, grows exponentially large in terms of n , and one cannot find the relevant fixed point by a plain iterative method. To resolve this problem, the BP iterations are equipped with a heuristic decimation process. This forms the basis of the BPGD algorithm.

B. BPGD algorithm

In this section we consider an instance of CCLDGM(l, r, L, w, n) ensemble. Equations (4) will be solved iteratively starting from the initial conditions $\eta_{i \rightarrow a}^{(0)} = \hat{\eta}_{a \rightarrow i}^{(0)} = 0$.

At iteration t the messages are $\eta_{i \rightarrow a}^{(t)}$ and $\hat{\eta}_{a \rightarrow i}^{(t)}$. We define the bias of a code-bit i at time t as

$$b_t(i) = \beta \sum_{a \in \partial i} \hat{\eta}_{a \rightarrow i}^{(t)}. \quad (5)$$

This represents the tendency of the code-bit to be 0 or 1. Our BPGD algorithm shown in Algorithm 1, uses a *decimation condition*. Let $\epsilon > 0$ a small quantity, $\alpha > 0$ a large quantity and T an iteration time. Typical values used in the simulations are $\epsilon = 0.01$, $\alpha = 4.25$ and $T = 10$. We say that the *decimation condition is fulfilled* if one of following occurs:

- i) After some time $\tau(\epsilon) < T$, the messages do not change significantly in two successive iterations, in the sense that $\frac{1}{nL} \sum_{(i,a) \in E(\Gamma)} |\hat{\eta}_{a \rightarrow i}^{(t)} - \hat{\eta}_{a \rightarrow i}^{(t-1)}| < \epsilon$ for $t > \tau(\epsilon)$.
- ii) For some i and $t < T$, $|b_t(i)| > \alpha$.
- iii) None of the above conditions has taken place $t \leq T$.

Algorithm 1 BPGD Algorithm

- 1) Begin with the graph instance $\Gamma^{(0)} \in \text{CCLDGM}(l, r, L, w, n)$.
 - 2) $t = 0$, Update equations (4) until the first time t_1 such that the decimation condition is fulfilled.
 - 3) Find $B = \max_i |b_{t_1}(i)|$.
 - 4) **If** $B = 0$, **then** randomly pick a code-bit i from range $[\frac{L-w}{2}, \frac{L+w}{2}]$ and fix it randomly to 0 or 1. **Else** pick a code-bit $i \in \{j \in C(\Gamma^{(k)}) \mid |b_{t_1}(j)| = B\}$ randomly and fix $u_i = \frac{1}{2}(\text{sign}(b_{t_1}(i)) + 1)$.
 - 5) Update $x_a^{(k+1)} = x_a^{(k)} - u_i$, for $a \in \partial i$, otherwise $x_a^{(k+1)} = x_a^{(k)}$. Then, update $\Gamma^{(k+1)} = \Gamma^{(k)} \setminus \{i\}$.
 - 6) **If** there exists an unfixed code-bit, **then** go to (2), **Else** finish and return \underline{u} .
-

It is not difficult to see that the complexity of the algorithm is $O(n^2 L^2)$. Note that after each decimation, rather than resetting messages to zero we continue with the previous messages.

An important element in the algorithm is that when there is no significant bias (step (4)) the random decimation occurs in a specific bounded interval centered around $L/2$ and of size w . We observe that this creates a seed from which the encoding process starts propagating through the ring. The usual hard decimation algorithms randomly choose (when there is no bias) a code-bit from the whole graph. We observed that when this prescription is used for CCLDGM instances, the distortion does not improve with respect to that of usual LDGM instances.

C. Optimal β and Relation to BSC

We have seen that a configuration \underline{u}^* maximizes the probability $\mu_{\beta}(\underline{u})$ for all $\beta > 0$. If μ_{β}^{BP} was an accurate approximation of μ_{β} , we would expect the maximum of $\mu_{\beta}^{\text{BP}}(\underline{u})$ to be independent of β . But μ_{β}^{BP} is not an exact description of μ_{β} . It is then natural to look at the performance of the BPGD procedure for different β and find an optimal parameter β_{opt} .

Figure 3 illustrates the performance of BPGD versus β for CCLDGM(3, 6, 64, 3, 2000) and LDGM(3, 6, 128000) ensembles. For this example, the first observation is that, for all

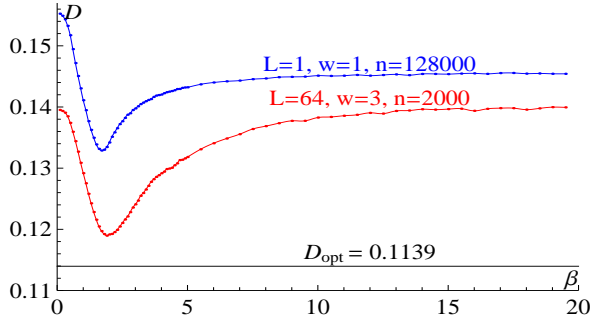


Fig. 3. BPGD distortion versus β for CCLDGM(3, 6, 64, 3, 2000) (bottom) and LDGM(3, 6, 128000) (top) ensembles.

$\beta > 0$, the coupled ensemble has smaller D_{BPGD} than the uncoupled ensemble. We also observe that each ensemble has an optimal parameter β_{opt} which lies between $\frac{3}{2}$ and $\frac{5}{2}$. In order to make comparison between coupled and uncoupled ensembles and as the distortion does not vary much when $\beta \in [\frac{3}{2}, \frac{5}{2}]$, we fix $\beta = 2$ in the rest of the paper. Table I shows that the BPGD distortion - for the value $\beta = 2$ - is still consistently higher than D_{opt} .

The measure μ_β also arises in the context of channel coding over a memoryless binary symmetric channel (BSC). Suppose that we use the LDGM code with factor graph Γ over a BSC with flipping probability p . Then the posterior probability that the un-coded message $\underline{u} \in \mathbb{F}_2^{nL/r}$ is sent given the received codeword $\underline{x} \in \mathbb{F}_2^n$ can be formally expressed just as in (3) with $\beta \rightarrow \ln\left(\frac{1-p}{p}\right)$. The difference between channel coding and lossy source coding lies in the distribution of \underline{X} . In lossy source coding, $\{X_a\}$ are n i.i.d. $\text{Ber}(1/2)$ variables; whereas in channel coding over a BSC, they are not in general i.i.d. and depend on the flipping probability p (or β).

V. DISTORTION SATURATION IN A CCLDGM ENSEMBLE

We provide simulation results which convincingly show that the BP distortion of a CCLDGM ensemble closely approaches the optimal distortion of the underlying LDGM ensemble as n , L and w grow large. To emphasize the effect of coupling, we consider the family of $(k, 2k)$ -regular LDGM ensembles. This family is known to yield a weak performance under the BPGD algorithm [12]. Indeed, although the optimal distortion is rather close to the rate-distortion bound and improves as k grows, the BPGD distortion becomes larger as k grows (see Table I). Analogous behaviors are well known to occur in channel coding [18].

Now consider a CCLDGM($k, 2k, L, w$) ensemble. Let D_z denote the average distortion of the reconstruction nodes lying at position z . Call D_z the *local distortion at position z* . Thus, the total average distortion is $D = \frac{1}{L} \sum_{z=0}^{L-1} D_z$. Call the vector $(D_0, D_1, \dots, D_{L-1})$ the *distortion profile*.

When we apply the BPGD algorithm, we observe a generic behaviour in the distortion profile. Figure 4 illustrates the distortion profile of CCLDGM($5, 10, L, w, n$) ensembles for different pairs of (L, w) and $n = 2000$. The profile is symmetric around $\frac{L}{2}$ (this expected in view of the algorithm). Consider the first $\frac{L}{2}$ components. In the range $[0, w - 1]$, the

TABLE II
SATURATED (BOLD) AND AVERAGE BPGD DISTORTION (PARENTHESES)
FOR CCLDGM(3, 6, L, w, n).

L	w	n		
		2000	4000	8000
32	2	0.1182 (0.1204)	0.1170 (0.1191)	0.1162 (0.1183)
	3	0.1171 (0.1205)	0.1161 (0.1195)	0.1157 (0.1190)
64	2	0.1185 (0.1195)	0.1172 (0.1183)	0.1164 (0.1174)
	3	0.1175 (0.1191)	0.1166 (0.1182)	0.1159 (0.1175)
128	2	0.1186 (0.1191)	0.1173 (0.1179)	
	3	0.1176 (0.1184)	0.1167 (0.1175)	

TABLE III
SATURATED (BOLD) AND AVERAGE BPGD DISTORTION (PARENTHESES)
FOR CCLDGM(4, 8, L, w, n).

L	w	n		
		2000	4000	8000
32	2	0.1165 (0.1205)	0.1148 (0.1188)	0.1134 (0.1175)
	3	0.1146 (0.1212)	0.1133 (0.1199)	0.1124 (0.1189)
	4	0.1135 (0.1226)	0.1125 (0.1215)	0.1118 (0.1208)
64	2	0.1170 (0.1190)	0.1151 (0.1171)	0.1138 (0.1157)
	3	0.1153 (0.1186)	0.1139 (0.1172)	0.1130 (0.1161)
	4	0.1144 (0.1189)	0.1133 (0.1177)	0.1125 (0.1169)
128	2	0.1172 (0.1182)	0.1153 (0.1162)	
	3	0.1156 (0.1172)	0.1141 (0.1157)	
	4	0.1148 (0.1170)	0.1136 (0.1158)	

local distortion is strictly decreasing. It starts at a value roughly equal to D_{BPGD} of the underlying ensemble and decreases to a value that we call the *saturation value*. Then, in the range $[w, \frac{L}{2}]$, the local distortions nearly remain constant and equal to the saturation value. We refer to the first interval and the second interval respectively as *the unsaturated part* and *the saturated part*. Therefore, the average distortion considerably decreases in the coupled ensembles.

Tables II, III and IV show the saturation value (in bold) and the average distortion (in parentheses) for $k = 3, 4, 5$ and different values of n , L and w . Each value is averaged over 100 random code instances and random source words. Inspection of the tables suggests that it approaches D_{opt} in the regime $n \gg L \gg w \gg 1$. The average distortion D_{BPGD} converges to the saturation value by increasing L . The reason is the unsaturated part essentially does not change for fixed k, w and n (see Figure 4), thus its contribution in the average distortion vanishes in the large L limit. As a result, we expect that D_{BPGD} gets very close to D_{opt} for a large enough $n \gg L \gg w \gg 1$ and the optimal β . Moreover, if k grows, D_{opt} converges to the rate-distortion bound. Thus, the D_{BPGD} of a regular CCLDGM ensemble can get close to the rate-distortion bound.

As mentioned before, the coupled chains used in channel coding are terminated by suitable boundary conditions and this incurs a rate loss of order $O(\frac{w}{L})$. In CCLDGM ensembles, we

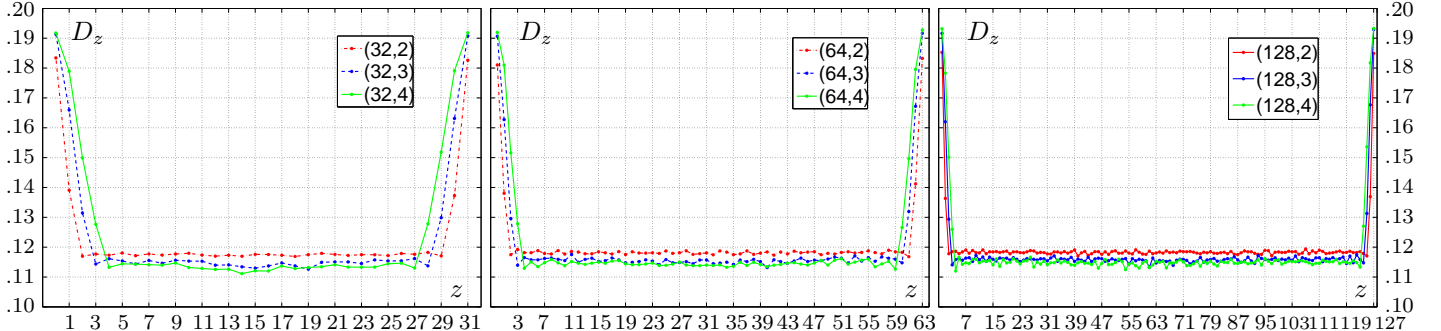


Fig. 4. BPGD distortion profiles for a periodic chain of $(5, 10, L, w, n)$ CCLDGM ensembles. The values of (L, w) are shown on the plots and $n = 2000$.

TABLE IV
SATURATED (BOLD) AND AVERAGE BPGD DISTORTION (PARENTHESES)
FOR CCLDGM $(5, 10, L, w, n)$.

L	w	n		
		2000	4000	8000
32	2	0.1175 (0.1229)	0.1153 (0.1207)	0.1138 (0.1192)
	3	0.1147 (0.1236)	0.1130 (0.1219)	0.1120 (0.1208)
	4	0.1134 (0.1256)	0.1120 (0.1242)	0.1111 (0.1234)
	5	0.1124 (0.1280)	0.1112 (0.1268)	0.1107 (0.1262)
64	2	0.1181 (0.1208)	0.1158 (0.1185)	0.1140 (0.1167)
	3	0.1155 (0.1199)	0.1138 (0.1182)	0.1125 (0.1169)
	4	0.1144 (0.1205)	0.1130 (0.1190)	0.1119 (0.1179)
	5	0.1137 (0.1213)	0.1124 (0.1200)	0.1116 (0.1192)
128	2	0.1182 (0.1196)	0.1158 (0.1172)	
	3	0.1158 (0.1180)	0.1140 (0.1161)	
	4	0.1148 (0.1178)	0.1133 (0.1162)	
	5	0.1141 (0.1179)	0.1128 (0.1166)	

do not pay this cost in the graph structure, but it manifests itself in the large local distortion in the $2w$ positions of the unsaturated part.

VI. CONCLUSION

We have observed that CCLDGM $(k, 2k, L, w)$ ensembles can asymptotically saturate the rate-distortion bound under a BPGD algorithm. Degrees $(k, 2k)$ have been chosen because it is known they lead to weaker results in the uncoupled case, but the saturation presumably occurs for a large class of irregular LDGM ensembles. As a result, using a right-regular CCLDGM ensemble with Poisson distribution on the code-bit nodes, would allow to achieve the rate-distortion bound for all rates $R \in (0, 1)$.

There are clearly a number of features of the algorithm that can be improved. Two important issues are its convergence rate, and its complexity. To address the first one, we can fix to zero the code-bits lying in range $[\frac{L-w}{2}, \frac{L+w}{2}]$ and then, we remove the reconstruction nodes (and source nodes) lying in ranges $[0, \frac{w}{2}]$ and $[L-1-\frac{w}{2}, L-1]$. The total rate of the code slightly increases but the convergence rate is improved. The complexity of the BPGD algorithm used here is $O(n^2 L^2)$. To reduce it one could use a “sliding window encoding” similar to the one used for coupled LDPC codes [3]. This would reduce

the complexity to $O(n^2 L)$. Finally the use of soft decimation techniques might also help improve the overall performance.

REFERENCES

- [1] A. J. Felstrom and K. S. Zigangirov, “Time-varying periodic convolutional codes with low density parity check matrix,” *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 2181–2190, 1999.
- [2] M. Lentmaier, A. Sridharan, D. J. Costello, Jr., and K. S. Zigangirov, “Iterative decoding threshold analysis for LDPC convolutional codes,” *To appear in IEEE Transactions on Information Theory*, 2008.
- [3] M. Lentmaier, A. Sridharan, K. S. Zigangirov, and D. J. Costello, “Terminated LDPC convolutional codes with thresholds close to capacity,” in *ISIT*, 2005, pp. 1372–1376.
- [4] M. Lentmaier, D. G. M. Mitchell, G. P. Fettweis, and D. J. Costello, “Asymptotically regular LDPC codes with linear distance growth and thresholds close to capacity,” in *Information Theory and Applications Workshop (ITA)*, January 2010, pp. 1–8.
- [5] S. Kudekar, T. J. Richardson, and R. L. Urbanke, “Threshold saturation via spatial coupling: why convolutional LDPC ensembles perform so well over the BEC,” <http://arxiv.org/abs/1001.1826>, 2009.
- [6] S. Kudekar, T. Richardson, and R. Urbanke, “Spatially Coupled Ensembles Universally Achieve Capacity under Belief Propagation,” *ArXiv*, Jan. 2012.
- [7] S. H. Hassani, N. Macris, and R. L. Urbanke, “Coupled graphical models and their thresholds,” in *Information Theory Workshop*, Dublin, 2010.
- [8] —, “Chains of mean field models,” *submitted to J. Stat. Mech: Theory and Experiment*, 2010.
- [9] —, “Threshold saturation in spatially coupled constraint satisfaction problems,” *ArXiv*, Dec 2011.
- [10] S. Kudekar and H. Pfister, “The effect of spatial coupling on compressive sensing,” in *48th Annual Allerton Conference*, Aug. 2010, pp. 347–353.
- [11] F. Krzakala, M. Mézard, F. Sausset, Y. Sun, and L. Zdeborová, “Statistical physics-based reconstruction in compressed sensing,” *ArXiv*, Sep. 2011.
- [12] M. Wainwright, E. Maneva, and E. Martinian, “Lossy source compression using low-density generator matrix codes: Analysis and algorithms,” *Information Theory, IEEE Transactions on*, vol. 56, no. 3, pp. 1351–1368, March 2010.
- [13] S. Ciliberti and M. Mzard, “The theoretical capacity of the parity source coder,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2005, no. 10, p. P10003, 2005.
- [14] M. Wainwright and E. Maneva, “Lossy source encoding via message-passing and decimation over generalized codewords of LDGM codes,” in *ISIT*, 2005, pp. 1493–1497.
- [15] T. Filler and J. Fridrich, “Binary quantization using belief propagation with decimation over factor graphs of ldgm codes,” in *Allerton*, 2007.
- [16] D. Castanheira and A. Gameiro, “Lossy source coding using belief propagation and soft-decimation over ldgm codes,” in *PIMRC*, sept. 2010, pp. 431–436.
- [17] V. Aref and R. L. Urbanke, “Universal rateless codes from coupled It codes,” in *Information Theory Workshop (ITW)*, oct. 2011, pp. 277–281.
- [18] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.
- [19] M. Mézard and A. Montanari, *Information, physics, and computation*, ser. Oxford graduate texts. Oxford University Press, 2009.